DOCUMENT RESUME

ED 370 980 TM 021 559

AUTHOR Roeber, Edward D.; And Others

TITLE Performance Assessment: A National Perspective. Policy Briefs Numbers 10 & 11. Special Double Issue.

INSTITUTION North Central Regional Educational Lab., Elmhurst, IL.
SPONS AGENCY Office of Educational Research and Improvement (ED), Washington, DC.
PUB DATE 90
CONTRACT 400-86-0004
NOTE 9p.
PUB TYPE Collected Works - Serials (022) -- Reports - Evaluative/Feasibility (142)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS *Educational Assessment; Educational Research; Educational Trends; Elementary Secondary Education; *National Competency Tests; Scoring; *State Programs; *Test Construction; Testing Problems; *Testing Programs; Training; Trend Analysis
IDENTIFIERS Indicators; Large Scale Programs; Michigan; *Performance Based Evaluation; Vermont

ABSTRACT

In the initial article by Edward Roeber, performance assessment is defined as an exercise in which the student demonstrates specific skills and competencies rather than selecting one of several predetermined answers to an exercise. Such as assessment contains four components: (1) a reason for the assessment; (2) a particular peformance to be evaluated; (3) exercises that elicit that performance; and (4) systematic rating procedures. Performance assessment, discussed from a national perspective, has emerged as a trend in itself because the stakes associated with large-scale assessment programs have increased so dramatically in recent years. Consequently, in a number of states, including Michigan, performance assessments are being developed as part of large-scale assessments. Performance assessment is needed for indicator systems at the national, state, and local levels. The effort required to develop, administer, and score performance assessments is large, but the rewards are well worth the effort. Some regional actions and agendas in the area of performance assessment are discussed, and the following guest commentaries relating to the feature article are presented: (1) "Performance Assessment: Living Up to Expectations" (Robert T. Linn); (2) "The Foundation of Performance Assessment: A Strong Training Program" (Richard J. Stiggins); and (3) "Performance Assessment in Vermont" (W. Ross Brewer). (SLD)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

Policy Briefs

Number 10 & 11, 1990: • *Special Double Issue*

A Publication of the North Central Regional Educational Laboratory

ED 370 980

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
☑ This document has been reproduced as received from the person or organization originating it.
☐ Minor changes have been made to improve reproduction quality.
• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

Performance Assessment

A National Perspective

by Edward D. Roeber, Michigan Department of Education

Editor's Note: *Due to the importance of testing and assessment issues in the NCREL region and nationwide, this Policy Brief is a special double issue. "Viewpoints" of state legislators are also included from their individual state perspective.*

Policy Briefs are reports on the status of current issues in education from a national perspective, descriptions of actions and agendas in the NCREL region, commentaries by experts from their particular point of view, and resources for further information.

What is performance assessment? Richard Stiggins (1987) defines performance assessment as an exercise in which the student demonstrates specific skills and competencies rather than selecting one of several predetermined answers to an exercise. This open-end demonstration can take place within everyday classroom activities or as a response to a carefully structured situation presented by a specially trained test administrator.

Stiggins indicates that such assessments contain four components: (1) a reason for the assessment, (2) a particular performance to be evaluated,(3) exercises that elicit that performance, and (4) systematic rating procedures. The response of the students may be given verbally, in writing, or in another manner (for example, singing) and may require simple or elaborate apparatus or none at all, with students working alone or in groups, spontaneously or rehearsed. The student's performance may be observed and scored on the spot, or may be recorded (on audio-or videotape) for later scoring.

Performance assessment has emerged as a trend in itself over the past few years because the stakes associated with the large-scale assessment programs have increased so dramatically. More states are initiating large-scale assessment programs; others are expanding the purposes of their programs. And some states have begun using assessment for competency testing purposes with results tied to student promotion or graduation. This is a significant change from the assessment programs begun 20 and 30 years ago which often were used to gather needs-assessment data on the educational system at state and local levels.

Additionally, assessment in specific subject areas has expanded beyond the traditional areas of reading and mathematics. Even definitions of these traditional areas are changing. For example, the area of mathematics has been redefined to incorporate topics such as conceptualization, problem solving, mental arithmetic, and the use of calculators and computers. Reading, rather than a static skill acquisition process, is now viewed as the interaction of the text and the reader with the ability to construct meaning influenced by prior knowledge about the topic, the metacognitive skills, and the reader's attitudes and self-perceptions.

Finally, large-scale national assessment programs are beginning to influence both assessment and instructional activities at the state and local levels. The United States is building stronger ties to assessment programs such as the International Educational Achievement Studies. These studies will continue to focus public attention on the quality of American schools.

TM021559

BEST COPY AVAILABLE

Even as external assessment programs exert greater influence on the instruction of students, another trend is occurring in the opposite direction: the recognition of the amount and type of informal assessment activities that go on in the classroom. Such informal assessments are a more important factor in improving individual student learning than formal external assessment programs. Several states have also begun the process of better teacher preparation for classroom assessment activities. Others have tried to build informal assessment models in such areas as reading, mathematics, and science that classroom teachers can use to complement and supplement the information provided by the formal, external assessments.

Within the context of large-scale assessment programs, there are several types of performance measures that have been used or proposed for use. Each of the four components described previously would have to be developed for each of the examples listed below. Although these examples do not include the more innovative informal performance measures under development in several states, they do illustrate the range and importance of performance assessment in large-scale programs. Following each subject area is a description of the measure: Art—draw, paint, participate in various art activities; Career Development—apply for a job, interview for a job, participate in group discussions; Employability Skills— participate in a work team to accomplish various work tasks, lead and follow in a work team; Mathematics—measure, use calculators and computers; Music—sing, dance, play a musical instrument; Reading—use reference materials, read in free time; Writing—write an essay or a letter, write for enjoyment; Speaking/Listening—speak in public, analyze conversations, communicate non-verbally, participate in a drama presentation.

These examples provide some of the most important outcomes that we hope our high school graduates take with them as they exit our schools. But what evidence do we have that they actually do have these skills? Classroom teachers may be able to gather evidence that students have learned these skills, both through observation in the course of classroom activities and the classroom-level examinations they use. However, are these skills assessed and reported in a manner comparable across districts that the public, at the local and state levels, can use to determine student performance? Probably not.

The result is that these important (some might argue the most important) outcomes of schools do not get the attention they should since the external pressure which large-scale assessment programs generate does not compel teachers to instruct these skills. Indeed, large-scale assessment programs may actually serve to reduce instructional time available for such activities since the pressure will be felt to teach the skills that are more easily testable using paper-and-pencil tests.

A decade ago it may have been important to assess such skills in order to have a sense of curricular completeness. Now the reasons are even more compelling. Many citizens are using the large-scale achievement results to judge the adequacy of our schools, and where they find them lacking, they are using the same results to spur schools to change. In addition, they use the results compiled over time to judge whether the changes have indeed been made.

Is it realistic to think in terms of assessing such skills within the context of large-scale assessment programs given that special measures requiring special test administration procedures, individual test administration, and special scoring processes will be needed? Is this feasible for programs that assess hundreds of thousands of students within a few weeks?

In Michigan and a growing number of other states, the answer is yes. It is feasible to develop such tests, administer them to at least statewide samples of students within a large-scale assessment program, and score and report student responses within the context of overall reports about student performance. New York has demonstrated that it is even feasible to administer performance tests in science on an every-pupil basis.

The bottom line is that performance assessment is needed for indicator systems at the national, state, and local levels. Indicator systems that use student achievement data will be used to determine student needs and encourage educators to meet those needs as well as to evaluate the efforts.

Because of the impact of assessment on the skills we want our teachers to teach and our students to achieve, it is vital that we include within the achievement measures used the most important skills that we want our students to accomplish. Performance assessment is needed to determine if students have achieved these skills. Such performance measures can be developed, administered, and reported for little cost. Although the effort required to develop and use such measures is considerable, the payoff is even greater. When such measures are used, they give the entire assessment an added aura of content validity. Since the heart of an indicator system is student outcome or achievement data, performance measures should be the heart of the achievement indicators used. Performance assessment is real, it's feasible, and now it's critical. ■

Edward Roeber is Supervisor of the Educational Assessment Program for the Michigan Department of Education and Co-Director of the Association of State Assessment Programs, an infor.nal technical assistance program for state testing.

Regional Action & Agendas

Illinois

The Illinois Goal Assessment Program (IGAP), established in 1985, mandates statewide assessment in six fundamental learning areas—language arts, reading, math, science, social sciences, and health and physical education—in grades 3, 6, 8, and 11. The purpose of the assessment is to measure the extent to which students have attained knowledge and skills with respect to specified goals in each learning area. In some learning areas, such as writing, measurement of skill is integral to the assessment and is well-defined. In others, such as fine arts, the IGAP assessment of skills is problematic due to the limitations of time, resources, and absence of agreed upon criteria or instruments. Within these limits, the Student Assessment Section attempts to tailor the assessment of each learning area in such a way as to maximize the opportunity for performance as well as knowledge assessment. Two examples:

Writing—The IGAP writing assessment requires each student to write an essay in response to one of several different types of prompts. Students are allotted 45 minutes for this task. All students at grades 3, 6, and 8—about 330,000—participated in April, 1990. These essays are scored according to criteria established by Illinois educators.

Science—The statewide IGAP science assessment, which begins in 1992, will not contain performance assessment items (manipulatives). Rather, using a portion of the $10 million grant monies which the Illinois legislature established for science and math literacy in Illinois, the State Board in 1990 funded a proposal by the Illinois Science Teachers Association to explore hands-on or process assessment instruments and protocols that can be made available for local schools to use in their local assessments. That grant project is underway. Two products will be a hands-on assessment video and a handbook to guide local teachers in performance assessment.

Indiana

Indiana's efforts in the area of performance assessment have officially taken off this year. All academic disciplines involved in the statewide testing program, (mathematics, language arts, science, and social studies) are considering ways to expand statewide assessment efforts in the area of performance assessment. First, the current test contract for Indiana Statewide Testing for Educational Progress (ISTEP) includes a provision to pilot performance testing in the areas of mathematics and social studies during the 1990-91 school year. Second, the 1990 General Assembly funded a Research and Development Center in the Department of Education, and one of the center's assignments will be to pilot innovative forms of assessment. The pilots will be quite small at this stage and are intended to offer the state information about the feasibility of conducting performance assessment on a statewide basis in the future.

Mathematics—The Indiana Department of Education plans to pilot performance testing in April, 1991 in grades 2, 3, and 6 through 8. The planned assessments were developed in compliance with the recently completed standards of the National Council of Teachers of Mathematics and the Indiana Curriculum Proficiency Guide in Mathematics. A small number of schools will be scheduled to pilot this performance test.

Language Arts—Indiana has had performance assessment in language arts for the past six years with its direct writing sample. Students in grades 3, 6, 8, 9, and 11 produce compositions under controlled conditions; the writing samples are scored holistically and analytically.

Plans for expanded performance assessment in language arts include improving and expanding the writing sample and adding proofreading/editing exercises and oral communication tasks.

In addition, 11 sites have been funded to develop pilot projects for portfolio assessment of writing and other language arts during the 1990-91 school year.

Science—The Science Proficiency Review/Revision Committee is revising the science proficiencies to bring them in line with the recommendations of the American Association for the Advancement for Science (AAAS) Project 2061 Report, *Science for All Americans*. It is developing assessment indicators for each desired learning outcome at the primary, upper elementary, middle school, and high school levels. In turn, the committee is recommending how each indicator can best be assessed, e.g., through multiple choice items, through other written means, or through performance of specified tasks.

Social Studies—A committee of social studies teachers is currently working with Indiana Department of Education staff to develop specifications for performance test items to be piloted in Spring, 1991 by a small number of volunteer schools. The goal of this effort is to develop items that will assess types of social studies learning that are not easily measured by other testing methods. Performance testing will engage students in specific activities such as map making, constructing time lines, interpretation of historical documents, and writing samples and group problem solving.

Iowa

The Director of Education has appointed a statewide committee composed of representatives from business, labor, and education to define "World Class Education" and to identify strategies for achieving it. The Committee has recommended that measurement of student performance be expanded to include a variety of behaviors to reflect what students are expected to learn. No specific performance areas have been identified for assessment.

Assessment plans will be formulated based upon Committee recommendations and projected costs. Oral and written tasks as well as multiple choice questions will be included.

A plan utilizing the results and a cost estimate for administering each segment of the assessment will be developed following the identification of behaviors to be assessed.

To reduce developmental costs, the Department will examine measures used in other states and work with staff from the Iowa Testing Program.

Michigan

Performance assessment has been a part of many of the assessment activities in the Michigan Educational Assessment Program (MEAP) over its 20-plus-year history. Beginning with a science performance assessment in 1974, such assess-

ments have been administered to statewide samples of students in the areas of mathematics, music, science, social studies, art, music, career development, and physical education/fitness. In most cases, the data are collected, scored, and interpreted by volunteers so that the cost of these special assessment projects is quite low. The Department was able to demonstrate the feasibility of such assessments.

Looking ahead, performance assessment figures prominently in future development plans for areas such as reading, mathematics, and science, as well as employability skills. Each area is currently assessed on an every-pupil basis at one or more grade levels, and performance assessment, either on all students or samples, is part of those plans. As assessments in each of these areas are developed or refined, performance assessments will be included among the assessments developed and used.

Minnesota

The Minnesota State Assessment Program through its Essential Learner Outcomes Assessments is committed to some form of performance-based assessment at each grade level and in each subject tested.

During the spring of 1990, a statewide sample of students in grades 6, 9, and 11 participated in a language arts assessment. The assessment included reading, writing, and proofreading/editing knowledge and skills. The reading assessment offered students complete poems, fiction stories, and/or current events articles and then presented multiple choice questions about the poem/story/article. Student writing samples were also collected at each grade level.

Statewide assessments to be administered in the spring of 1991 include science at grades 6, 9, and 11; and health at grades 4, 8 and 11. Performance based "lab" stations will be part of the science assessment at each grade. Students will be observed performing a variety of tasks and evaluated according to an established scoring criteria. Performance-based experiences will also be included in the health assessment.

In addition to the mandatory statewide assessment program, the Minnesota Test Itembank (MIDEBANK) offers its users an ever-expanding array of performance-based evaluation opportunities. Music and art items calling for actual performance by students are accompanied by scoring criteria. Similar performance-based experiences are available in science and other subjects. As the bank grows and evolves, performance-based assessment, surveys, and open-ended items are a major emphasis.

Ohio

The following are some of the many Ohio activities related to student performance assessment. In July 1990, the State Board of Education announced its intent to move toward performance-based graduation requirements for all students. Many vocational education and special education programs have been performance-based. Since 1984, all districts have been implementing competency-based education programs including student performance assessment in reading, writing, and mathematics. Currently the state is funding "Classroom of the Future" projects and providing technical assistance related to performance assessment, including the use of portfolios. In November 1990, the first statewide graduation test included two student samples.

Reform legislation enacted in 1989 requires the State Board of Education to identify excellent and deficient schools and districts using indicators that include student performance measures. The same bill also provides for public school districts to grant equivalent high school credit to eligible adults who are able to demonstrate competencies equivalent to those acquired through high school courses. The State Board is studying a number of issues related to implementing such performance-based programs.

Thoughtful discussion will continue regarding the design of an integrated assessment system that yields results useful to students, educators, parents, policymakers, prospective employers, and other stakeholders in a manner that is legally, technically, and fiscally defensible.

Wisconsin

Wisconsin continues to have a strong interest in the topic of student performance assessment and is monitoring local, state, and federal developmental activities. Other than two state-sample writing assessments conducted during the 1980s, there have been no state-level activities of this type.

The Department of Public Instruction is discussing performance assessment issues and possibilities as part of the 1992-93 biennial budget development process. In addition, the Governor's Commission on Schools for the 21st Century, which will develop recommendations for the biennial budget, is considering the topic.

The Department believes performance assessment has many positive qualities, especially as designed and implemented at the classroom, school, and district levels. There are, however, significant issues of reliability, validity, cost, and effort that must be carefully examined prior to statewide implementation within a large-scale assessment program. The work underway in several states is of interest to Wisconsin and will be watched closely to determine the extent to which such assessment activities could be integrated into a statewide program.

It will not be until June 1991 that new policy and budget initiatives will be determined for the next two fiscal years. ■

Guest Commentary

Performance Assessment: Living Up to Expectations

by Robert L. Linn, University of Colorado at Boulder

Assessment policies and procedures at the local, state, and national levels are in a period of rapid transformation. The demands for greater accountability and higher standards are increasing the salience of assessment at the same time that there is a growing dissatisfaction with traditional methods of assessment and increasing concern about the unintended negative consequences of high-stakes testing programs. The resulting ferment has set the stage for major changes in assessment policies and practice.

Performance assessments, such as those described by Ed Roeber (this issue), provide the vision that is guiding many of the efforts to transform assessment. Possibly the most important promise of performance assessments is that they will facilitate improvements in instruction and learning. Performance assessments are designed to engage students in solving problems and performing substantial tasks of importance in their own right. Such assessments are expected to correspond more closely to important instructional goals than the multiple-choice items found on a standardized test. Indeed, the major promise of performance assessments is that they are so congruent with educational goals and good instructional practice that they will enhance instruction and facilitate the achievement of important educational goals.

If performance assessments realize their promise and lead to better educational consequences, they certainly will be worth the substantial effort that the construction and implementation of such assessments will require. Of course, performance assessments are not all alike. They will not all be equally valid measures of important learning outcomes nor will they all have equally salutary effects on learning and instruction. Too often there is a great gap between intentions and implementation. Hence, there is a need to be able to determine not only the degree to which expectations are realized but also the relative validity and the nature of the consequences, both intended and unintended, of different performance assessments. In short, we need criteria for evaluating performance assessments.

Criteria for Evaluating Performance Assessments

Traditional criteria, the most important of which are validity and reliability, do exist for judging measurements. These criteria, although relevant to performance assessments, must be thought of in broader terms than they normally are when applied to standardized measures. In *Complex, Performance-Based Assessment: Expectations and Validation Criteria* (see Resource list), Eva Baker, Stephen Dunbar, and I describe ways in which the criteria need to be expanded and applied to the evaluation of performance assessments. In the limited space available here, only three of the criteria we propose will be summarized.

Consequences

The concept of validity has been expanded in the past few years to include concerns about the intended and unintended consequences of measurement. This expanded notion is particularly relevant to assessments, either traditional or direct performance assessments, that are intended to have an impact on student learning. The major promise needs to be verified. Evidence must be gathered on a systematic basis regarding both the unintended and intended effects of assessments on the ways in which teachers and students spend their time. It is not enough to assume that assessments will have the intended effects just because they have greater face validity. Such assumptions should be checked, and ways of increasing the likelihood of achieving intended outcomes must be identified.

Generalizability and Transfer

It is obvious in the case of a multiple-choice test that there is a need to be concerned about the degree to which results generalize to other ways of demonstrating knowledge and understanding. The concerns also apply to performance assessments, however. We need to know the degree to which performance on a written essay, for example, generalizes to performance on other writing tasks. The same can be said for a performance involving a scientific experiment. Do problem-solving skills demonstrated on one lab problem transfer to other problems?

Fairness

Fairness is clearly an issue for any assessment but is a particular concern for high-stakes assessments. Because performance assessments often require substantial amounts of student time, it is usually impractical to administer more than a few problems. Indeed, in some cases a performance assessment may consist of only a single assignment (e.g., a laboratory experiment). Insensitivity of assessments to differences in background knowledge and the experiences of students outside of school can lead to unfairness. It is critical that such effects be minimized if performance assessments are to be used in making decisions about individual students or schools.

Consequences, generalizability, and fairness are only a few of the criteria to be considered in evaluating performance assessments. Seeking evidence related to these and other criteria will be critical if performance assessments are to realize their promise and not simply come and go as another educational fad. Hard evidence will be needed not only to satisfy skeptics and justify the higher costs of performance assessments but also to distinguish between effective and ineffective assessments. ■

Robert L. Linn is Professor of Education and Co-Director of the Center for Research on Evaluation, Standards, and Student Testing at the University of Colorado at Boulder.

Linn, R.L., Baker, E.L., & Dunbar, S.B. (1990). Complex, Performance-Based Assessment: Expectations and Validation Criteria. Los Angeles: UCLA Center for Research on Evaluation, Standards, and Student Testing.

Guest Commentary

The Foundation of Performance Assessment: A Strong Training Program

by Richard J. Stiggins, Northwest Regional Educational Laboratory

In one of the assessment workshops I do with and for teachers on the meaning of high-quality classroom assessment, one of the first activities is a brainstorming session in which participants list all the student characteristics, attributes, or traits that they assess daily in their classrooms. Invariably, these lists grow very long.

Then, as we examine the list together, I ask them to contemplate which and how many of the important educational outcomes listed can effectively be translated into objective paper-and-pencil test items. Again invariably, we are able to generate many outcomes that can be assessed in this way. Obviously, paper-and-pencil assessments can play a key role in documenting educational outcomes.

However, the other inference that becomes painfully obvious in this exercise is that there is a broad array of outcomes—including many of those we value most—that cannot be assessed via objective paper-and-pencil tests. Other modes of assessment must be used if we are to cover those key achievement targets. The two additional modes we talk about are observation and professional judgment (or performance assessment) and personal communication with students.

One key to quality assessment, I explain to teachers, is to have in mind a clear vision of the achievement target so we can select an assessment mode that fits. A second key is to know how to use each of the available assessment methods well. Each mode carries with it the potential of sound or unsound assessment. The trick is to know the difference. The third and final key is to know how to marry each valued target to appropriate methods.

Acknowledging that no single method can serve all our needs, what kinds of achievement targets are best reflected in performance assessments? There are many, as illustrated in the examples presented throughout this *Policy Brief*. They fall into two categories: targets that manifest themselves in achievement-related behaviors exhibited by the student (i.e., communication skills, psychomotor skills, and behaviors indicative of affective states) and targets that take the form of complex achievement-related products that students create (i.e., written reports, art and craft products and the like).

Performance assessments require that the assessor: (1) observe the behavior as exhibited or examine the product that is reflective of achievement, and (2) apply clearly articulated performance criteria so as to make a sound professional judgment regarding the level of proficiency demonstrated. In sound performance assessments, the underlying rigor is apparent, i.e., the assessor has defined the target, elicited the proper performance-articulated standards against which to compare each student's work, and generated careful records of student performance. Through these steps, assessors build objectivity into subjective assessments.

Note that we are not talking about "intuitions" about students here. Words such as "impressions" of performance, having a "sense" or a "feeling" of how well kids did, or making "informal" observations are not synonymous with, nor are they a part of, sound performance assessments.

Like paper-and-pencil tests, sound performance assessments must adhere to certain rules of evidence. While the rules are not the same for observation and judgment-based assessments as they are for paper-and-pencil tests, they are every bit as important. Those who adhere to the rules of evidence develop assessments that produce quality results. Those who violate the rules risk harming students and perpetuating the myth that the only truly "objective" (read: fair, unbiased) assessments are those that rely on paper-and-pencil item formats.

The key to success lies in understanding how to use performance assessment methodology—or any assessment methodology for that matter—effectively and efficiently. This takes training. It is tempting for us to ascribe responsibility for mastering that assessment wisdom to "those measurement experts," thus absolving ourselves of responsibility for knowing about assessment.

But the large-scale assessment examples that permeate this *Policy Brief*—visible, expensive, and politically important as they are—represent only a small fraction of one percent of the assessments conducted in schools. The other 99.9 percent are designed, developed, and used by teachers. Classroom assessments provide nearly all of the performance-related information that teachers, parents, and students themselves use in determining how students will spend their academic lives.

The critical assessment problem we face in the 1990s is that the vast majority of educators completed programs of professional preparation (graduate and undergraduate, preservice and inservice) that included virtually no training in assessment. The newly emerging emphasis on performance assessment methods holds great promise for expanding the array of important student outcomes we can assess. But it will only reach its potential if all users know how to use it well—as one of many assessment tools they have at their disposal.

As a result of a decade of research on performance assessment, we now understand what teachers and administrators need to know about them if they are to use them effectively. Only one step remains: as we think about the exciting and very promising future of performance assessment and paper-and-pencil assessment and assessment based on personal communications, we must: (a) constantly regard the alignment between and among these methods and valued achievement targets, and (b) allocate both assessment and professional development resources to provide both teachers and their supervisors with the opportunity to develop the competence and wisdom they need to assess effectively and to take full advantage of assessment results. ■

Richard J. Stiggins is the Director of the Center for Classroom Assessment at the Northwest Regional Educational Laboratory in Portland, Oregon.

Northwest Regional Educational Laboratory. Hitting the target: Classroom assessment workshops help take the fear out of testing. (Includes a list of eight available workshops and ten video-based presentations.) Available from: Richard Stiggins, Center for Classroom Assessment, Northwest Regional Educational Laboratory, 101 S.W. Main, Suite 500, Portland, Oregon 97204. Phone: 800-547-6339.

Guest Commentary

Performance Assessment in Vermont

by W. Ross Brewer, Vermont Department of Education

Beginning this year, Vermont will assess the achievement of fourth and eighth grade students in writing and mathematics using three methods: a portfolio, a "best piece," and a uniform test. In the coming years, the assessment system will be expanded to the high school level and will include portfolios, best pieces, and uniform assessments in science, history, the social sciences, and the arts. The results of the assessments will be reported by each school at School Report Day and by the state in the *Condition of Education Report.* Descriptions of the assessment methods used are as follows:

Portfolio—Each fourth and eighth grade student will keep portfolios of his or her work in writing and mathematics. The portfolio will contain the student's significant class work from the course of the year. Because the work in the portfolio is the result of regular classroom work, the teacher will already have assessed each piece individually. Teachers will also review portfolio work using the criteria developed by the mathematics and writing committees. Teachers will be able to compare their judgments with assessments made by outside evaluators.

Best Piece—At the end of the year, students and teachers will review the portfolio. The students will select from their portfolios the works they feel represent their best efforts for the year and will write an explanation of their choices.

Uniform Test—All fourth and eighth grade students will take a test that uses equivalent tasks administered under the same conditions.

Evaluation—Student work will be assessed by two trained Vermont teachers. Where there are significant disagreements, additional teachers will evaluate the portfolios.

Vermont has adopted this approach to assessing student performance because it fits with the state's traditions while emphasizing the knowledge and skills that we know are the key to our future. It builds on the professionalism of our teachers by assessing student performance using methods that have integrity. School Report Day carries on the Vermont tradition of town meetings to discuss important public issues. Because there are multiple indicators that present a rich array of information about student achievement, the Vermont system takes the emphasis off achievement of "a number" and at the same time provides parents, policymakers, educators, and taxpayers information that is important to them. It allows for the fair comparison of achievement in schools while discouraging the simplistic ranking which is so popular—and destructive.

The Vermont approach brings with it a unique set of challenges. Because what we are doing is new and very different, credibility will be an important issue. The work that is being evaluated is much more complex than that which is usually assessed using multiple choice or fill-in-the-blank tests. The challenge is characterized at one level by a legislator who asked, "Aren't kids supposed to know fractions; isn't that what we should be testing for?"

With the use of the new assessment techniques, people will not be able to read results in ranks, percentiles, grade levels, or stanines. When school-level results are reported, there will be many who challenge the results, noting the new and untried technology. Most importantly, teachers who feel that they too are being assessed will be on the firing line during a time when they are implementing a new, more complex, curriculum.

We have not survived these challenges yet, but we do have a strategy for dealing with them. First, we work from the bottom up. Educators developed our approach after listening to hundreds of Vermonters—educators, board members, business leaders, legislators, and representatives of the media—in forums and hearings and ultimately gained their vocal support.

Secondly, professional development is stressed and 40 percent of our assessment budget is earmarked for assessment training. Teachers' committees oversee the professional development program and design and run the bulk of the workshops. Writing and mathematics teachers were also responsible for designing the portfolio program.

Finally, in the area of scoring and reporting, Vermont's approach is to have clearly articulated, widely discussed standards of performance and a well-trained group of teachers who will apply these standards to students' work. Having two teachers evaluate each portfolio will begin to address the issue of subjectivity that is bound to arise.

The challenge to revamp Vermont's assessment procedures and reporting practices will, of course, lead to controversy. Mistakes will be made, but our guiding philosophy is to listen when they are pointed out to us and make adjustments in our policy and practices whenever and where ever needed. ■

W. Ross Brewer is the Director of Planning and Policy Development in the Vermont Department of Education.

Viewpoints

NCREL asked legislators in each state to respond to the following question: *What are the good points and problems, if any, of the current performance assessment policy in your state?*

Art Ollie, Chair, Iowa House Education Committee, 413 Ruth Place, Clinton, Iowa 52732:

"Iowa policy mandates that schools, upon advice from an advisory board including parents, adopt short- and long-term goals including student achievement goals along with a plan for assessing the accomplishment of achievement. They must report results to citizens and to the Department of Education."

"Experience with the assessment mandate is too brief (effective 7/1/89) to evaluate. The policy provides more flexibility than uniformity, so there is no statewide consistency in student assessment. Also, I am concerned about the extensive use of standardized tests at the K-3 level."

"Iowa school districts tailor assessment to fit the needs of their objectives for students and the nature of the communities in which the students live. I believe that, in general, Iowa assessment practices are for the purposes of improving instruction and benefiting the students assessed. The integrity of assessment in Iowa schools is considered very high."

Ken Nelson, Chair, Education Finance Division, State Representative, 367 State Office Building, St. Paul, Minnesota 55155:

"Our state assessment program is limited primarily to testing. We need to develop broader methods of assessing student achievement and progress that can reflect individual student learning. Those methods might include student projects and portfolios, interview, conferences, observations, and evaluations."

"Our state assessment program is aligned with the state learner outcomes so that curriculum, assessment, and instruction work together. This means that assessment, curriculum, and instruction all work towards the same goal."

Cooper Snyder, Chairman, Senate Education Committee, Ohio Senate, Statehouse, 1st Floor, Columbus, Ohio 43216:

"Our performance assessment program is in an embryonic stage. We are just beginning implementation. Ohio's approach is to combine data collection into one streamlined system that provides management information that will allow policymakers to really get a handle on cause and effect. I hope to be able to be more responsive a year from now." ■

State Contacts

ILLINOIS
Illinois State Board of Education
100 North First Street, E-230
Springfield, IL. 62777-0001
Tom Kerins: 217/782-4823

INDIANA
Indiana Department of Education
Room 229, State House
Indianapolis, Indiana 46204-2798
Linda Bond: 317/232-6610

IOWA
Iowa Department of Education
Grimes State Office Building
Des Moines, Iowa 50319-0146
Max Morrison: 515/281-5274

MICHIGAN
Department of Education
Educational Assessment Program
P.O. Box 30008
Lansing, Michigan 48909
Ed Roeber: 517/373-8393

MINNESOTA
Minnesota Department of Education
Capitol Square Building
550 Cedar Street
St. Paul, Minnesota 55101
Jim Colwell: 616/296-5119

OHIO
Ohio Department of Education
Division of Educational Services
65 S. Front Street, Room 811
Columbus, Ohio 43266-0308
Roger Trent: 614/466-3224

WISCONSIN
Department of Public Instruction
125 South Webster
P.O. Box 7841
Madison, Wisconsin 53707
Tom Stefonek: 608/266-1782

...nions expressed in the commentaries do not necessarily reflect the views of the NCREL staff or Board. Facts and ideas presented in NCREL's Policy Briefs are intended to survey a current issue and not to advocate a particular position.

Policy Briefs
A publication of the North Central
Regional Educational Laboratory
295 Emroy Avenue
Elmhurst, Illinois 60126

Telephone: (708) 941-7677
FAX: (708) 941-7163
...TE: ncrel.lab

...Nowakowski, *Executive Director*
...cy Fulford, *Editor*
...rianne Kroeger, *Production Editor*
...ue Casiello, *Desktop Publishing Coord.*

...his publication is based on work sponsored wholly or in part by the Office of Educational Research and Improvement (OERI), Department of Education under Contract Number 400-86-0004. The content of this publication does not necessarily reflect the views of OERI, the Department of Education, or any other agency in the U.S. Government.

Nonprofit Organization
U.S. Postage PAID
Elmhurst, Illinois
Permit No. 338